

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Opperman, Hermann
Ozkaynak, Engin
Rueger, David C.
Kuberasanpath, Thangavel
- (ii) TITLE OF INVENTION: Osteogenic Proteins
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESSES:
 - (A) ADDRESSEE: Lahive & Cockfield
 - (B) STREET: 60 State Street
 - (C) CITY: Boston
 - (D) STATE: Massachusetts
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 02109
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette, 3.5 inch,
720kb storage
 - (B) COMPUTER: IBM XT
 - (C) OPERATING SYSTEM: DOS 3.30
 - (D) SOFTWARE: Word Perfect 5.0
- (vi) CURRENT APPLICATION DATA:
 - (B) FILING DATE: 18-Oct-90
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 569,920
 - (B) FILING DATE: 20-Aug-90
 - (C) APPLICATION NUMBER: US 315,342
 - (D) FILING DATE: 23-Feb-89
 - (E) APPLICATION NUMBER: US 422,699
 - (F) FILING DATE: 17-Oct-89

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 139 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala	Ala	Arg	Pro	Leu	Lys	Arg	Arg	Gln
1				5				
Pro	Lys	Lys	Thr	Asn	Glu	Leu	Pro	His
10					15			
Pro	Asn	Lys	Leu	Pro	Gly	Ile	Phe	Asp
	20					25		
Asp	Gly	His	Gly	Ser	Arg	Gly	Arg	Glu
		30					35	
Val	Cys	Arg	Arg	His	Glu	Leu	Tyr	Val
			40					45
Arg	Phe	Arg	Asp	Leu	Gly	Trp	Leu	Asp
				50				
Trp	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
55					60			
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ala
	65					70		
Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
		75					80	
Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
			85					90
Val	His	Leu	Met	Lys	Pro	Asp	Val	Val
				95				
Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
100					105			
Leu	Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr
	110					115		

Asp	Ser	Ser	Asn	Asn	Val	Ile	Leu	Arg
		120					125	
Lys	His	Arg	Asn	Met	Val	Val	Lys	Ala
			130					135
Cys	Gly	Cys	His.					

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1930 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Bovinae

(F) TISSUE TYPE: bone

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: mouse embryo

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGAATTCCGC	TGCCAGGCAC	AGGTGCGCCG	TCTGGTCCTC	40
CCCGTCTGGC	GTCAGCCGAG	CCCGACCAGC	TACCAGTGGA	80
TGCGCGCCGG	CTGAAAGTCC	GAG ATG GCT ATG CGT		115
		Met Ala Met Arg		
		1		
CCC GGG CCA CTC TGG CTA TTG GGC CTT GCT CTG				148
Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu				
5	10		15	
TGC GCG CTG GGA GGC GGC CAC GGT CCC GGT CCC				181
Cys Ala Leu Gly Gly Gly His Gly Pro Gly Pro				
	20		25	

CCG	CAC	ACC	TGT	CCC	CAG	CGT	CGC	CTG	GGA	GCG	214
Pro	His	Thr	Cys	Pro	Gln	Arg	Arg	Leu	Gly	Ala	
			30						35		
CGC	GAC	CGG	GAC	ATG	CAG	CGT	GAA	ATC	CTG	CCG	247
Arg	Asp	Arg	Asp	Met	Gln	Arg	Glu	Ile	Leu	Pro	
		40					45				
GTG	CTC	GGG	CTA	CCG	GGA	CGC	CCC	GAC	CCC	GTG	280
Val	Leu	Gly	Leu	Pro	Gly	Arg	Pro	Asp	Pro	Val	
	50					55					
CAC	AAC	CCG	CCG	CTG	CCC	GGC	ACG	CAG	CGT	GCG	313
His	Asn	Pro	Pro	Leu	Pro	Gly	Thr	Gln	Arg	Ala	
60					65					70	
CCC	CTC	TTC	ATG	TTG	GAC	CTA	TAC	CAC	GCC	ATG	346
Pro	Leu	Phe	Met	Leu	Asp	Leu	Tyr	His	Ala	Met	
			75						80		
ACC	GAT	GAC	GAC	GAC	GGC	GGG	CCA	CCA	CAG	GCT	379
Thr	Asp	Asp	Asp	Asp	Gly	Gly	Pro	Pro	Gln	Ala	
			85					90			
CAC	TTA	GGC	CGT	GCC	GAC	CTG	GTC	ATG	AGC	TTC	412
His	Leu	Gly	Arg	Ala	Asp	Leu	Val	Met	Ser	Phe	
		95					100				
GTC	AAC	ATG	GTG	GAA	CGC	GAC	CGT	ACC	CTG	GGC	445
Val	Asn	Met	Val	Glu	Arg	Asp	Arg	Thr	Leu	Gly	
	105					110					
TAC	CAG	GAG	CCA	CAC	TGG	AAG	GAA	TTC	CAC	TTT	478
Tyr	Gln	Glu	Pro	His	Trp	Lys	Glu	Phe	His	Phe	
115					120					125	
GAC	CTA	ACC	CAG	ATC	CCT	GCT	GGC	GAG	GCT	GTC	511
Asp	Leu	Thr	Gln	Ile	Pro	Ala	Gly	Glu	Ala	Val	
				130					135		
ACA	GCT	GCT	GAG	TTC	CGG	ATC	TAC	AAA	GAA	CCC	544
Thr	Ala	Ala	Glu	Phe	Arg	Ile	Tyr	Lys	Glu	Pro	
			140						145		

AGC	ACC	CAC	CCG	CTC	AAC	ACA	ACC	CTC	CAC	ATC	577
Ser	Thr	His	Pro	Leu	Asn	Thr	Thr	Leu	His	Ile	
		150						155			
AGC	ATG	TTC	GAA	GTG	GTC	CAA	GAG	CAC	TCC	AAC	610
Ser	Met	Phe	Glu	Val	Val	Gln	Glu	His	Ser	Asn	
	160					165					
AGG	GAG	TCT	GAC	TTG	TTC	TTT	TTG	GAT	CTT	CAG	643
Arg	Glu	Ser	Asp	Leu	Phe	Phe	Leu	Asp	Leu	Gln	
170					175					180	
ACG	CTC	CGA	TCT	GGG	GAC	GAG	GGC	TGG	CTG	GTG	676
Thr	Leu	Arg	Ser	Gly	Asp	Glu	Gly	Trp	Leu	Val	
				185					190		
CTG	GAC	ATC	ACA	GCA	GCC	AGT	GAC	CGA	TGG	CTG	709
Leu	Asp	Ile	Thr	Ala	Ala	Ser	Asp	Arg	Trp	Leu	
		195						200			
CTG	AAC	CAT	CAC	AAG	GAC	CTG	GGA	CTC	CGC	CTC	742
Leu	Asn	His	His	Lys	Asp	Leu	Gly	Leu	Arg	Leu	
		205					210				
TAT	GTG	GAA	ACC	GCG	GAT	GGG	CAC	AGC	ATG	GAT	775
Tyr	Val	Glu	Thr	Ala	Asp	Gly	His	Ser	Met	Asp	
	215					220					
CCT	GGC	CTG	GCT	GGT	CTG	CTT	GGA	CGA	CAA	GCA	808
Pro	Gly	Leu	Ala	Gly	Leu	Leu	Gly	Arg	Gln	Ala	
225					230					235	
CCA	CGC	TCC	AGA	CAG	CCT	TTC	ATG	GTA	ACC	TTC	841
Pro	Arg	Ser	Arg	Gln	Pro	Phe	Met	Val	Thr	Phe	
				240					245		
TTC	AGG	GCC	AGC	CAG	AGT	CCT	GTG	CGG	GCC	CCT	874
Phe	Arg	Ala	Ser	Gln	Ser	Pro	Val	Arg	Ala	Pro	
			250					255			
CGG	GCA	GCG	AGA	CCA	CTG	AAG	AGG	AGG	CAG	CCA	907
Arg	Ala	Ala	Arg	Pro	Leu	Lys	Arg	Arg	Gln	Pro	
		260					265				

AAG	AAA	ACG	AAC	GAG	CTT	CCG	CAC	CCC	AAC	AAA	940
Lys	Lys	Thr	Asn	Glu	Leu	Pro	His	Pro	Asn	Lys	
	270					275					
CTC	CCA	GGG	ATC	TTT	GAT	GAT	GGC	CAC	GGT	TCC	973
Leu	Pro	Gly	Ile	Phe	Asp	Asp	Gly	His	Gly	Ser	
280					285					290	
CGC	GGC	AGA	GAG	GTT	TGC	CGC	AGG	CAT	GAG	CTC	1006
Arg	Gly	Arg	Glu	Val	Cys	Arg	Arg	His	Glu	Leu	
				295					300		
TAC	GTC	AGA	TTC	CGT	GAC	CTT	GGC	TGG	CTG	GAC	1039
Tyr	Val	Arg	Phe	Arg	Asp	Leu	Gly	Trp	Leu	Asp	
			305					310			
TGG	GTC	ATC	GCC	CCC	CAG	GGC	TAC	TCT	GCC	TAT	1072
Trp	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser	Ala	Tyr	
		315					320				
TAC	TGT	GAG	GGG	GAG	TGT	GCT	TTC	CCA	CTG	GAC	1105
Tyr	Cys	Glu	Gly	Glu	Cys	Ala	Phe	Pro	Leu	Asp	
	325					330					
TCC	TGT	ATG	AAC	GCC	ACC	AAC	CAT	GCC	ATC	TTG	1138
Ser	Cys	Met	Asn	Ala	Thr	Asn	His	Ala	Ile	Leu	
335					340					345	
CAG	TCT	CTG	GTG	CAC	CTG	ATG	AAC	CCA	GAT	GTT	1171
Gln	Ser	Leu	Val	His	Leu	Met	Lys	Pro	Asp	Val	
				350					355		
GTC	CCC	AAG	GCA	TGC	TGT	GCA	CCC	ACC	AAA	CTG	1204
Val	Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys	Leu	
			360					365			
AGT	GCC	ACC	TCT	GTG	CTG	TAC	TAT	GAC	AGC	AGC	1237
Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr	Asp	Ser	Ser	
		370					375				
AAC	AAT	GTC	ATC	CTG	CGT	AAA	CAC	CGT	AAC	ATG	1270
Asn	Asn	Val	Ile	Leu	Arg	Lys	His	Arg	Asn	Met	
	380					385					

GTG GTC AAG GCC TGT GGC TGC CAC	1294
Val Val Lys Ala Cys Gly Cys His	
390 400	

TGAGGCCCCG CCCAGCATCC TGCTTCTACT ACCTTACCAT	1334
CTGGCCGGGC CCCTCTCCAG AGGCAGAAAC CCTTCTATGT	1374
TATCATAGCT CAGACAGGGG CAATGGGAGG CCCTTCACTT	1414
CCCCTGGCCA CTTCTGTCTA AAATTCTGGT CTTTCCCAGT	1454
TCCTCTGTCC TTCATGGGGT TCGGGGGCTA TCACCCCGCC	1494
CTCTCCATCC TCCTACCCCA AGCATAGACT GAATGCACAC	1534
AGCATCCCAG AGCTATGCTA ACTGAGAGGT CTGGGGTCAG	1574
CACTGAAGGC CCACATGAGG AAGACTGATC CTTGGCCATC	1614
CTCAGCCCAC AATGGCAAAT TCTGGATGGT CTAAGAAGCC	1654
CTGGAATTCT AAACATAGATG ATCTGGGCTC TCTGCACCAT	1694
TCATTGTGGC AGTTGGGACA TTTTATAGGTA TAACAGACAC	1734
ATACACTTAG ATCAATGCAT CGCTGTACTC CTTGAAATCA	1774
GAGCTAGCTT GTTAGAAAAA GAATCAGAGC CAGGTATAGC	1814
GGTGCATGTC ATTAATCCCA GCGCTAAAGA GACAGAGACA	1854
GGAGAATCTC TGTGAGTTCA AGGCCACATA GAAAGAGCCT	1894
GTCTCGGGAG CAGGAAAAAA AAAAAAACG GAATTC	1930

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 139 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala	Val	Arg	Pro	Leu	Arg	Arg	Arg	Gln
1				5				
Pro	Lys	Lys	Ser	Asn	Glu	Ileu	Pro	Gln
10					15			

Ala	Asn	Arg	Leu	Pro	Gly	Ile	Phe	Asp
	20					25		
Asp	Val	His	Gly	Ser	His	Gly	Arg	Gln
		30					35	
Val	Cys	Arg	Arg	His	Glu	Leu	Tyr	Val
			40					45
Ser	Phe	Gln	Asp	Leu	Gly	Trp	Leu	Asp
				50				
Trp	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
55					60			
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser
	65					70		
Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
		75					80	
Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
			85					90
Val	His	Leu	Met	Lys	Pro	Asn	Ala	Val
				95				
Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
100					105			
Leu	Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr
	110					115		
Asp	Ser	Ser	Asn	Asn	Val	Ile	Leu	Arg
		120					125	
Lys	Ala	Arg	Asn	Met	Val	Val	Lys	Ala
			130					135
Cys	Gly	Cys	His.					

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1956 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: no
(iv) ANTI-SENSE: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bovinae
(F) TISSUE TYPE: bone
(vii) IMMEDIATE SOURCE:
(A) LIBRARY: human hippocampus
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGAATTCCGG	CCACAGTGGC	GCCGGCAGAG	CAGGAGTGGC	40
TGGAGGAGCT	GTGGTTGGAG	CAGGAGGTGG	CACGGCAGGG	80
CTGGAGGGCT	CCCTATGAGT	GGCGGAGACG	GCCCAGGAGG	120
CGCTGGAGCA	ACAGCTCCCA	CACCGCACCA	AGCGGTGGCT	160
GCAGGAGCTC	GCCCATCGCC	CCTGCGCTGC	TCGGACCGCG	200
GCCACAGCCG	GACTGGCGGG	TACGGCGGCC	ACAGAGGCAT	240
TGGCCGAGAG	TCCCAGTCCG	CAGAGTAGCC	CCGGCCTCGA	280
GGCGGTGGCG	TCCCGGTCCT	CTCCGTCCAG	GAGCCAGGAC	320
AGGTGTCGCG	CGGCGGGGCT	CCAGGGACCG	CGCCTGAGGC	360
CGGCTGCCCCG	CCCGTCCCCG	CCCGCCCCCG	CGCCCGCCGC	400
CCGCCGAGCC	CAGCCTCCTT	GCCGTCGGGG	CGTCCCCAGG	440
CCCTGGGTCTG	GCCGCGGAGC	CGATGCGCGC	CCGCTGAGCG	480
CCCCAGCTGA	GCGCCCCCGG	CCTGCC ATG	ACC GCG CTC	518

Met Thr Ala Leu

1

CCC	GGC	CCG	CTC	TGG	CTC	CTG	GGC	CTG	GCG	CTA	551
Pro	Gly	Pro	Leu	Trp	Leu	Leu	Gly	Leu	Ala	Leu	
5				10					15		
TGC	GCG	CTG	GGC	GGG	GGC	GGC	CCC	GGC	CTG	CGA	584
Cys	Ala	Leu	Gly	Gly	Gly	Gly	Pro	Gly	Leu	Arg	
			20						25		

CCC	CCG	CCC	GGC	TGT	CCC	CAG	CGA	CGT	CTG	GGC	617
Pro	Pro	Pro	Gly	Cys	Pro	Gly	arg	Arg	Leu	Gly	
			30						35		
GCG	CGC	GAC	CGG	GAC	GTG	CAG	CGC	GAG	ATC	CTG	650
Ala	Arg	Asp	Arg	Asp	Val	Gln	Arg	Glu	Ile	Leu	
		40					45				
GCG	GTG	CTC	GGG	CTG	CCT	GGG	CGG	CCC	CGG	CCC	683
Ala	Val	Leu	Gly	Leu	Pro	Gly	Arg	Pro	Arg	Pro	
	50					55					
CGC	GCG	CCA	CCC	GCC	GCC	TCC	CGG	CTG	CCC	GCG	716
Arg	Ala	Pro	Pro	Ala	Ala	Ser	Arg	Leu	Pro	Ala	
60					65					70	
TCC	GCG	CCG	CTC	TTC	ATG	CTG	GAC	CTG	TAC	CAC	749
Ser	Ala	Pro	Leu	Phe	Met	Leu	Asp	Leu	Tyr	His	
				75					80		
CGC	ATG	GCC	GGC	GAC	GAC	GAC	GAG	GAC	GGC	GCC	782
Arg	Met	Ala	Gly	Asp	Asp	Asp	Glu	Asp	Gly	Ala	
			85					90			
GCG	GAG	GCC	CTG	GGC	CGC	GCC	GAC	CTG	GTC	ATG	815
Ala	Glu	Ala	Leu	Gly	Arg	Ala	Asp	Leu	Val	Met	
		95					100				
AGC	TTC	GTT	AAC	ATG	GTG	GAG	CGA	GAC	CGT	GCC	848
Ser	Phe	Val	Asn	Met	Val	Glu	Arg	Asp	Arg	Ala	
	105					110					
CTG	GGC	CAC	CAG	GAG	CCC	CAT	TGG	AAG	GAG	TTC	881
Leu	Gly	His	Gln	Glu	Pro	His	Trp	Lys	Glu	Phe	
115					120					125	
CGC	TTT	GAC	CTG	ACC	CAG	ATC	CCG	GCT	GGG	GAG	914
Arg	Phe	Asp	Leu	Thr	Gln	Ile	Pro	Ala	Gly	Glu	
				130					135		
GCG	GTC	ACA	GCT	GCG	GAG	TTC	CGG	ATT	TAC	AAG	947
Ala	Val	Thr	Ala	Ala	Glu	Phe	Arg	Ile	Tyr	Lys	
			140					145			

GTG	CCC	AGC	ATC	CAC	CTG	CTC	AAC	AGG	ACC	CTC	980
Val	Pro	Ser	Ile	His	Leu	Leu	Asn	Arg	Thr	Leu	
		150					155				
CAC	GTC	AGC	ATG	TTC	CAG	GTG	GTC	CAG	GAG	CAG	1013
His	Val	Ser	Met	Phe	Gln	Val	Val	Gln	Glu	Gln	
		160				165					
TCC	AAC	AGG	GAG	TCT	GAC	TTG	TTC	TTT	TTG	GAT	1046
Ser	Asn	Arg	Glu	Ser	Asp	Leu	Phe	Phe	Leu	Asp	
		170			175					180	
CTT	CAG	ACG	CTC	CGA	GCT	GGA	GAC	GAG	GGC	TGG	1079
Leu	Gln	Thr	Leu	Arg	Ala	Gly	Asp	Glu	Gly	Trp	
				185						190	
CTG	GTG	CTG	GAT	GTC	ACA	GCA	GCC	AGT	GAC	TGC	1112
Leu	Val	Leu	Asp	Val	Thr	Ala	Ala	Ser	Asp	Cys	
			195					200			
TGG	TTG	CTG	AAG	CGT	CAC	AAG	GAC	CTG	GGA	CTC	1145
Trp	Leu	Leu	Lys	Arg	His	Lys	Asp	Leu	Gly	Leu	
		205					210				
CGC	CTC	TAT	GTG	GAG	ACT	GAG	GAC	GGG	CAC	AGC	1178
Arg	Leu	Tyr	Val	Glu	Thr	Glu	Asp	Gly	His	Ser	
		215				220					
GTG	GAT	CCT	GGC	CTG	GCC	GGC	CTC	CTG	GGT	CAA	1211
Val	Asp	Pro	Gly	Leu	Ala	Gly	Leu	Leu	Gly	Gln	
		225			230					235	
CGG	GCC	CCA	CGC	TCC	CAA	CAG	CCT	TTC	GTG	GTC	1244
Arg	Ala	Pro	Arg	Ser	Gln	Gln	Pro	Phe	Val	Val	
				240						245	
ACT	TTC	TTC	AGG	GCC	AGT	CCG	AGT	CCC	ATC	CGC	1277
Thr	Phe	Phe	Arg	Ala	Ser	Pro	Ser	Pro	Ile	Arg	
			250					255			
ACC	CCT	CGG	GCA	GTG	AGG	CCA	CTG	AGG	AGG	AGG	1310
Thr	Pro	Arg	Ala	Val	Arg	Pro	Leu	Arg	Arg	Arg	
		260					265				

CAG	CCG	AAG	AAA	AGC	AAC	GAG	CTG	CCG	CAG	GCC	1343
Gln	Pro	Lys	Lys	Ser	Asn	Glu	Leu	Pro	Gln	Ala	
270						275					
AAC	CGA	CTC	CCA	GGG	ATC	TTT	GAT	GAC	GTC	CAC	1376
Asn	Arg	Leu	Pro	Gly	Ile	Phe	Asp	Asp	Val	His	
280					285					290	
GGC	TCC	CAC	GGC	CGG	CAG	GTC	TGC	CGT	CGG	CAC	1409
Gly	Ser	His	Gly	Arg	Gln	Val	Cys	Arg	Arg	His	
				295					300		
GAG	CTC	TAC	GTC	AGC	TTC	CAG	GAC	CTC	GGC	TGG	1442
Glu	Leu	Tyr	Val	Ser	Phe	Gln	Asp	Leu	Gly	Trp	
			305						310		
CTG	GAC	TGG	GTC	ATC	GCT	CCC	CAA	GGC	TAC	TCG	1475
Leu	Asp	Trp	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser	
		315					320				
GCC	TAT	TAC	TGT	GAG	GGG	GAG	TGC	TCC	TTC	CCA	1508
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser	Phe	Pro	
	325					330					
CTG	GAC	TCC	TGC	ATG	AAT	GCC	ACC	AAC	CAC	GCC	1541
Leu	Asp	Ser	Cys	Met	Asn	Ala	Thr	Asn	His	Ala	
335					340					345	
ATC	CTG	CAG	TCC	CTG	GTG	CAC	CTC	ATG	AAG	CCA	1574
Ile	Leu	Gln	Ser	Leu	Val	His	Leu	Met	Lys	Pro	
				350					355		
AAC	GCA	GTC	CCC	AAG	GCG	TGC	TGT	GCA	CCC	ACC	1607
Asn	Ala	Val	Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	
			360						365		
AAG	CTG	AGC	GCC	ACC	TCT	GTG	CTC	TAC	TAT	GAC	1640
Lys	Leu	Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr	Asp	
			370				375				
AGC	AGC	AAC	AAC	GTC	ATC	CTG	CGC	AAA	GCC	CGC	1673
Ser	Ser	Asn	Asn	Val	Ile	Leu	Arg	Lys	Ala	Arg	
				380		385					

AAC ATG GTG GTC AAG GCC TGC GGC TGC CAC	1703
Asn Met Val Val Lys Ala Cys Gly Cys His	
390	395
TGAGTCAGCC CGCCCAGCCC TACTGCAGCA ATTCACTGGC	1743
CGTCGTTTTTA CAACGTGTGA CTGGGAAAAC CCTGGCGTTA	1783
CCCAACTTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG	1823
CTGGCTAATA GCGAAGAGGC CCCGCACCGA TCGCCCTTCC	1863
CAACAGTTGC GCCCCAGTGA ATGGCGAATG GCAAATTGTA	1903
AGCGTTAATA TTTTGTTAAA ATTCGCGTTA AATTTTTT	1941

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Cys Xaa ₁ Xaa ₂ His Glu Leu Tyr Val Xaa ₃ Phe	
1	5 10
Xaa ₄ Asp Leu Gly Trp Xaa ₅ Asp Trp Xaa ₆ Ile	
	15 20
Ala Pro Xaa ₇ Gly Tyr Xaa ₈ Ala Tyr Tyr Cys	
	25 30
Glu Gly Cys Xaa ₉ Phe Pro Leu Xaa ₁₀ Ser Xaa ₁₁	
	35 40
Met Asn Ala Thr Asn His Ala Ile Xaa ₁₂ Thr	
	45 50
Leu Xaa ₁₃ Xaa ₁₄ Xaa ₁₅ Xaa ₁₆ Xaa ₁₇ Xaa ₁₈ Val	
	55
Pro Lys Xaa ₁₉ Cys Cys Ala Pro Thr Xaa ₂₀ Leu	
60	65

Xaa₂₁ Ala Xaa₂₂ Ser Val Leu Tyr Xaa₂₃ Asp
70 75
Xaa₂₄ Ser Xaa₂₅ Asn Val Xaa₂₆ Leu Xaa₂₇ Lys
80 85
Xaa₂₈ Pro Asn Met Val Val Xaa₂₉ Ala Cys Gly
90 95
Cys His,

wherein Xaa₁ = (Lys or Arg); Xaa₂ = (Lys or Arg);
Xaa₃ = (Ser or Arg); Xaa₄ = (Arg or Gln); Xaa₅ =
(Gln or Leu); Xaa₆ = (Ile or Val); Xaa₇ = (Glu or
Gln); Xaa₈ = (Ala or Ser); Xaa₉ = (Ala or Ser);
Xaa₁₀ = (Asn or Asp); Xaa₁₁ = (Tyr or Cys); Xaa₁₂ =
(Val or Leu); Xaa₁₃ = (His or Asn); Xaa₁₄ = (Phe or
Leu); Xaa₁₅ = (Ile or Met); Xaa₁₆ = (Asn or Lys);
Xaa₁₇ = (Glu, Asp or Asn); Xaa₁₈ = (Thr, Ala or
Val); Xaa₁₉ = (Pro or Ala); Xaa₂₀ = (Gln or Lys);
Xaa₂₁ = (Asn or Ser); Xaa₂₂ = (Ile or Thr); Xaa₂₃ =
(Phe or Tyr); Xaa₂₄ = (Asp, Glu or Ser); Xaa₂₅ =
(Ser or Asn); Xaa₂₆ = (Ile or Asp); Xaa₂₇ = (Lys or
Arg); Xaa₂₈ = (Tyr, Ala or His); and Xaa₂₉ = (Arg
or Lys).

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 437 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:

2

								TCC	ACG	GGG	9
								Ser	Thr	Gly	
								1			
AGC	AAA	CAG	CGC	AGC	CAG	AAC	CGC	TCC	AAG	ACG	42
Ser	Lys	Gln	Arg	Ser	Gln	Asn	Arg	Ser	Lys	Thr	
	5					10					
CCC	AAG	AAC	CAG	GAA	GCC	CTG	CGG	ATG	GCC	AAC	75
Pro	Lys	Asn	Gln	Glu	Ala	Leu	Arg	Met	Ala	Asn	
15					20					25	
GTG	GCA	GAG	AAC	AGC	AGC	AGC	GAC	CAG	AGG	CAG	108
Val	Ala	Glu	Asn	Ser	Ser	Ser	Asp	Gln	Arg	Gln	
				30					35		
GCC	TGT	AAG	AAG	CAC	GAG	CTG	TAT	GTC	AGC	TTC	141
Ala	Cys	Lys	Lys	His	Glu	Leu	Tyr	Val	Ser	Phe	
			40					45			
CGA	GAC	CTG	GGC	TGG	CAG	GAC	TGG	ATC	ATC	GCG	174
Arg	Asp	Leu	Gly	Trp	Gln	Asp	Trp	Ile	Ile	Ala	
		50					55				
CCT	GAA	GGC	TAC	GCC	GCC	TAC	TAC	TGT	GAG	GGG	207
Pro	Glu	Gly	Tyr	Ala	Ala	Tyr	Tyr	Cys	Glu	Gly	
	60					65					
GAG	TGT	GCC	TTC	CCT	CTG	AAC	TCC	TAC	ATG	AAC	240
Glu	Cys	Ala	Phe	Pro	Leu	Asn	Ser	Tyr	Met	Asn	
70					75					80	
GCC	ACC	AAC	CAC	GCC	ATC	GTG	CAG	ACG	CTG	GTC	273
Ala	Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	
				85					90		

CAC	TTC	ATC	AAC	CCG	GAA	ACG	GTG	CCC	AAG	CCC	306
His	Phe	Ile	Asn	Pro	Glu	Thr	Val	Pro	Lys	Pro	
			95						100		
TGC	TGT	GCG	CCC	ACG	CAG	CTC	AAT	GCC	ATC	TCC	339
Cys	Cys	Ala	Pro	Thr	Gln	Leu	Asn	Ala	Ile	Ser	
		105					110				
GTC	CTC	TAC	TTC	GAT	GAC	AGC	TCC	AAC	GTC	ATC	372
Val	Leu	Tyr	Phe	Asp	Asp	Ser	Ser	Asn	Val	Ile	
	115					120					
CTG	AAG	AAA	TAC	AGA	AAC	ATG	GTG	GTC	CGG	GCC	405
Leu	Lys	Lys	Tyr	Arg	Asn	Met	Val	Val	Arg	Ala	
125					130					135	
TGT	GGC	TGC	CAC	TAG	CTC	TCC	GAG	AAT	CAG		437
Cys	Gly	Cys	His								

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1					5					10	
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			15						20		
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa
			25				30				
Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa
			35			40					
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
45						50					55


```

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
                        60                      65
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                        70                      75
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                        80                      85
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
                        90                      95
Xaa Cys Xaa,
100

```

wherein each Xaa independently represents one of the 20 naturally occurring L-isomer, α -amino acids.

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 97 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1                        5                      10
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                        15                      20
Xaa Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa
                        25                      30
Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                        35                      40
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
45                        50                      55

```

```

Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Xaa Xaa Xaa
      60                      65
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
      70                      75
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
      80                      85
Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa,
      90                      95

```

wherein each Xaa independently represents one of the 20 naturally occurring L-isomer, α -amino acids.

- (2) INFORMATION FOR SEQ ID NO:9:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 136 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

          Pro  Leu  Arg  Arg  Arg  Gln
            1              5
Pro  Lys  Lys  Ser  Asn  Glu  Leu  Pro  Gln
            10              15
Ala  Asn  Arg  Leu  Pro  Gly  Ile  Phe  Asp
            20
Asp  Val  Asn  Gly  Ser  His  Gly  Arg  Gln
 25              30
Val  Cys  Arg  Arg  His  Glu  Leu  Tyr  Val
    35              40
Ser  Phe  Gln  Asp  Leu  Gly  Trp  Leu  Asp
            45              50

```

Tyr	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
			55					60
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser
			65					
Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
70					75			
Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
	80					85		
Val	His	Leu	Met	Lys	Pro	Asn	Ala	Val
		90					95	
Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
			100					105
Leu	Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr
			110					
Asp	Glu	Ser	Asn	Asn	Val	Ile	Leu	Arg
115					120			
Lys	Ala	Arg	Asn	Met	Val	Val	Lys	Ala
	125					130		
Cys	Gly	Cys	His.					
		135						

- (2) INFORMATION FOR SEQ ID NO:10:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 133 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

						Arg	Arg	Gln
						1		
Pro	Lys	Lys	Ser	Asn	Glu	Leu	Pro	Gln
	5					10		

Ala	Asn	Arg	Leu	Pro	Gly	Ile	Phe	Asp
		15					20	
Asp	Val	Asn	Gly	Ser	His	Gly	Arg	Gln
			25					30
Val	Cys	Arg	Arg	His	Glu	Leu	Tyr	Val
				35				
Ser	Phe	Gln	Asp	Leu	Gly	Trp	Leu	Asp
40					45			
Tyr	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
	50					55		
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser
		60					65	
Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
			70					75
Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
				80				
Val	His	Leu	Met	Lys	Pro	Asn	Ala	Val
85					90			
Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
	95					100		
Leu	Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr
		105					110	
Asp	Glu	Ser	Asn	Asn	Val	Ile	Leu	Arg
			115					120
Lys	Ala	Arg	Asn	Met	Val	Val	Lys	Ala
				125				
Cys	Gly	Cys	His.					
130								

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 170 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

						Ser	Gln	Gln
						1		
Pro	Phe	Val	Val	Thr	Phe	Phe	Arg	Ala
	5					10		
Ser	Pro	Ser	Pro	Ile	Arg	Thr	Pro	Arg
		15					20	
Ala	Val	Arg	Pro	Leu	Arg	Arg	Arg	Gln
			25					30
Pro	Lys	Lys	Ser	Asn	Glu	Leu	Pro	Gln
				35				
Ala	Asn	Arg	Leu	Pro	Gly	Ile	Phe	Asp
40					45			
Asp	Val	Asn	Gly	Ser	His	Gly	Arg	Gln
	50					55		
Val	Cys	Arg	Arg	His	Glu	Leu	Tyr	Val
		60					65	
Ser	Phe	Gln	Asp	Leu	Gly	Trp	Leu	Asp
			70					75
Tyr	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
				80				
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser
85					90			
Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
	95					100		
Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
		105					110	
Val	His	Leu	Met	Lys	Pro	Asn	Ala	Val
			115					120
Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
				125				

Leu 130	Ser	Ala	Thr	Ser	Val 135	Leu	Tyr	Tyr
Asp	Glu 140	Ser	Asn	Asn	Val	Ile 145	Leu	Arg
Lys	Ala	Arg 150	Asn	Met	Val	Val	Lys 165	Ala
Cys	Gly	Cys	His. 170					